

# 期末專案參考資訊



# 物件辨識



# 資料集座標讀取

- Kitti資料集中label的格式如下:

```
Misc 0.00 0 -1.82 804.79 167.34 995.43 327.94 1.63 1.48 2.37 3.23 1.59 8.55 -1.47
```

- 其中:
  - (804.79 ,167.34 ,995.43 ,327.94)分別為( $X_{min}$ ,  $Y_{min}$ ,  $X_{max}$ ,  $Y_{max}$ )
- YOLO的label格式如下:
  - 1 0.716797 0.395833 0.216406 0.147222
  - 分別為<object-class> <x> <y> <width> <height>
  - <x> <y> <width> <height>的數值範圍為(0.0, 1.0]，此範圍是YOLO訓練所需格式



# 資料集座標讀取

- 因此可以自己寫程式，讀取txt檔，將 $(X_{min}, Y_{min}, X_{max}, Y_{max})$ 轉為`<object-class> <x> <y> <width> <height>`，這樣就不必手動標記了，並儲存成與圖片相同的檔名
  - `object-class` → 依自己label數量而定有n種，就有0~n-1，可以只label一種就好，方便計算iou、準確率等等
  - $x = f(X_{min})$
  - $y = f(Y_{min})$
  - $width = f(X_{max} - X_{min})$
  - $height = f(Y_{max} - Y_{min})$
  - $f(x)$ 代表將x座標根據圖片長或寬正規化到(0.0, 1.0]



# 計算準確度

- ground truth 就是你標記的座標(從kitty轉成YOLO的那個座標)。  
預測結果座標要和這個座標比對準確率。
- YOLO除了輸出辨識結果圖片之外也可以從code中取得預測框的座標，可以使用iou來算準確率。
  - 一個預測框跟所有ground truth的iou=0就是代表誤框(或沒框到)。
  - 一個預測框跟ground truth的iou可以設定0.5來決定是否為誤框(因為可能框不完整)。若有多個iou就以最大的iou為主。



# YOLO參數

- YOLO參數說明，請參考：
  - <https://chih-sheng-huang821.medium.com/%E6%B7%B1%E5%BA%A6%E5%AD%B8%E7%BF%92-%E7%89%A9%E4%BB%B6%E5%81%B5%E6%B8%ACyolov1-yolov2%E5%92%8Cyolov3-cfg-%E6%AA%94%E8%A7%A3%E8%AE%80-75793cd61a01>
- YOLOv1~v4不同版本只是網路架構有稍微修改而已，其餘參數皆大同小異。



# 軌跡預測



# 軌跡預測

- 可以自行決定要使用Social LSTM, Social GAN等方法，請根據Lecture內容中提到的參數去調整來訓練。
- 軌跡預測，就說若是時間不足，可以用LSTM不一定要social LSTM，但social準確度會更高。
- 其中ADE(平均位移誤差 Average displacement error)與FDE(終點位移誤差 Final displacement error)，可以在進行validate後輸出的結果看到數值結果。